Gli indicatori bibliometrici: riflessioni sparse per un uso attento e consapevole

Tessa Piazzini

Quando negli anni Cinquanta del secolo scorso Eugene Garfield mise a punto il Citation Index (Garfield 1955, 2006; Garfield, Sher e Torpie 1964), con il duplice scopo di aiutare i ricercatori a selezionare gli articoli di maggior interesse per il proprio campo di ricerca e i bibliotecari ad attuare delle efficaci politiche di acquisizione, probabilmente non avrebbe mai immaginato l'impatto che tale strumento avrebbe avuto nella comunità scientifica.

Da sempre abituato a confrontare e discutere le proprie idee,[1] il mondo scientifico e accademico si trovò di fronte ad un nuovo parametro di valutazione, l'Impact factor, che, seppure per riflesso, avrebbe finito per influire, e non poco, sulle modalità stesse di circolazione del sapere, soprattutto in certi ambiti disciplinari.

Dopo l'Impact factor, o meglio, probabilmente in seguito alla valenza distorta con cui si cominciò a utilizzarlo[2], l'ambiente scien-

[1] Una buona sintesi delle tappe storiche dell'evoluzione del concetto di *peer review* si trova in (Di Donato 2010, cap. 1, in particolare p. 13-28),.

[2]Lo stesso Garfield aveva prudenzialmente fatto notare che however, analyses based on citation counts must be challenged with the question, What is the relationship between citation frequency and the historical impact or importance of the work

JLIS.it. Vol. 1, n. 1 (Giugno/June 2010), p. 63–86.
DOI: 10.4403/jlis.it-24

tifico ha continuato a proporre altri indicatori che nel corso degli anni sono andati ad implementare una griglia per la valutazione quantitativa[3], continuamente in divenire.

Pur nella consapevolezza che la componente prioritaria del giudizio di merito sul prodotto di una ricerca scientifica è e debba rimanere la valutazione qualitativa (sia essa codificata nella formula del peer reviewing sia attuata con modalità meno rigidamente strutturate, garantite da canali di pubblicazione alternativi all'editoria commerciale tradizionale[4]), non si può non considerare che l'oggettivazione numerica degli indicatori bibliometrici sia sentita comunque come la prova numericamente tangibile della validità di una ricerca[5].

Tale visione pare essere condivisa anche da parte di chi di ricerca, del suo finanziamento, produzione e valutazione si occupa[6]:

---

cited?' High citation counts reflect impact but may or may not reflect intrinsic worth. The data obtained from citation analysis are always relative rather than absolute. (Garfield, Sher e Torpie 1964).

[3] Valutazione quantitativa che si articola sostanzialmente in due componenti: l'analisi citazionale e l'analisi di utilizzo.

[4]Rientrano in questo settore soprattutto tutte quelle applicazioni del cosiddetto web 2.0 che permettono una circolazione più rapida del sapere ed una condivisione di giudizi e opinioni più diffusa grazie, ad esempio, alla possibilità di inserire un commento o di attivare una discussione attorno ad un articolo pubblicato secondo la filosofia Open Access. Anche i modelli ibridi, come riviste online Open Access ma peer reviewed e dotate di Impact factor, si possono annoverare in questo gruppo.

[5]Questo è vero soprattutto per le comunità scientifiche del settore STM, mentre per le scienze umane l'applicazione della bibliometria è ancora molto difficoltosa e poco codificata; la causa è da ricercare probabilmente nel modello di circolazione del sapere umanistico" basato principalmente sulla pubblicazione di monografie, riflesso di una tempistica di ricerca più lenta e dilatata nel tempo, rispetto alla produzione delle scienze dure, che vede nell'articolo di rivista la propria concretizzazione più tipica.

[6]Si vedano ad esempio le *Linee guida per la valutazione della ricerca* predisposte dal CIVR nel 2006 (CIVR 2006), ma anche la posizione del Ministero dell'Università e della ricerca scientifica così come espressa dal Decreto Ministeriale 28 luglio 2009 prot. n. 89/2009 *Valutazione dei titoli e delle pubblicazioni scientifiche.*

ministeri, università, enti pubblici e privati, etc.

Attualmente, in Italia[7] e altrove[8], la valutazione della ricerca è sempre più strettamente collegata non tanto, e non solo, alla garanzia di circolazione, diffusione e condivisione dei risultati ottenuti, quanto alla ripartizione degli esigui fondi per la ricerca stessa, alle modalità di reclutamento di docenti e ricercatori in ambito accademico, alla capacità di attrarre capitali da finanziatori pubblici e privati[9].

Qualunque bando di ricerca richiede infatti al candidato di produrre un'attestazione di merito, che si concretizza nella produzione di una bibliografia dei propri contributi migliori; qualunque docente o ricercatore universitario è tenuto a rendere conto al ministero stesso, tramite le anagrafi della ricerca nazionali o locali[10], della

(http://www.miur.it/0006Menu_C/0012Docume/0015Atti_M/7921Valuta_cf2.htm).

[7] Nel panorama italiano la valutazione della ricerca è stata recentemente oggetto di varie azioni normative: dai precedenti organismi CNVSU e CIVR nel 2008 è stata istituita l'ANVUR (Agenzia nazionale di valutazione del sistema universitario e di ricerca), il cui regolamento di organizzazione e funzionamento è stato approvato dal Presidente della Repubblica solo a febbraio 2010 con il D.P.R. 01/02/2010: *Regolamento concernente la struttura ed il funzionamento dell'Agenzia nazionale di valutazione del sistema universitario e della ricerca, ANVUR, a norma dell'articolo 2, comma 140, del decreto-legge 3 ottobre 2006, n. 262, convertito, con modificazioni, dalla legge 24 novembre 2006, n. 286*. (In attesa di pubblicazione sulla Gazzetta Ufficiale). Per un'efficace sintesi del complesso iter normativo si veda Antonella De Robbio, *L'Open Access come strategia per la valutazione delle produzioni intellettuali* (De Robbio 2010).

[8] A livello europeo manca però un'omogeneità dei criteri di valutazione, carenza che diventa un problema nel caso di bandi di finanziamento internazionali perché rende difficile la comparabilità della validità dei progetti presentati.

[9] Anche tra gli interventi correttivi proposti nelle Linee guida del CIVR «favorire la diffusione e la circolazione dei risultati della ricerca» si colloca solamente al terzo posto dopo «migliorare il collegamento istituzionale tra i risultati della valutazione, selezione dei programmi e allocazione delle risorse» e «avviare procedure sistematiche di raffronto tra organismi di ricerca nazionali e internazionali» (CIVR 2006).

[10] La situazione italiana attuale è abbastanza disomogenea a causa della non obbligatorietà di registrazione ad un'anagrafe nazionale di ricerca, di fatto formalmente

propria attività e del proprio ruolo in seno all'università di appartenenza, presentando i propri lavori scientifici: in entrambi i casi viene sempre richiesta l'indicazione degli indici citazionali.

Pertanto la conoscenza degli indicatori bibliometrici e degli strumenti per il loro recupero sta assumendo sempre più un peso finanziario" ed economico che va ben oltre il riconoscimento del proprio valore e merito da parte dei membri della comunità scientifica di riferimento.

Spesso però è proprio sul reperimento di tali dati che si addensano le maggiori difficoltà, sia da parte degli utenti sia da parte dei bibliotecari, perché se è vero che ormai Impact factor, Citation index e il già consolidato H-index sono noti alla maggioranza degli appartenenti ai due gruppi, lo stesso non si può dire per le modalità di recupero, per le banche dati e gli strumenti da utilizzare, per i limiti e le caratteristiche di fiognuno di questi indicatori, che, se usati correttamente, possono costituire un ausilio importante nel processo di valutazione. A volte non viene in soccorso neanche la normativa esistente o le indicazioni fornite dai bandi, non sempre

---

non esistente.

chiare o complete[11].

Lo scopo di questo articolo è quindi quello di fornire, senza alcuna pretesa di completezza, spunti di riflessione e consigli pratici per un utilizzo consapevole degli indicatori bibliometrici più consolidati, segnalandone i limiti e le peculiarità, senza entrare nel dibattito sulla loro applicazione nel mondo Open Access[12] e sulla necessità di creare e utilizzare altri parametri quantitativi di valutazione[13] che

[11]Si veda ad esempio il Decreto Ministeriale 28 luglio 2009 prot. n. 89/2009 *Valutazione dei titoli e delle pubblicazioni scientifiche*, cit., in particolare il comma 4 dell'art. 3: «Nell'ambito dei settori scientifico-disciplinari in cui ne è riconosciuto l'uso a livello internazionale, le Commissioni nel valutare le pubblicazioni si avvalgono anche dei seguenti indici:

- numero totale delle citazioni;
- numero medio di citazioni per pubblicazione;
- impact factor" totale;
- impact factor" medio per pubblicazione;
- combinazioni dei precedenti parametri atte a valorizzare l'impatto della produzione scientifica del candidato (indice di Hirsch o simili).»

Fa riflettere la presenza di un indicatore (impact factor totale) la cui esistenza non è attestata in letteratura, così come la generalizzazione indice di Hirsch e simili"; nessun cenno infine viene fatto all'esigenza di indicare la banca dati da cui sono stati tratti gli indici.

[12] Nel ricco panorama di letteratura sull'argomento si vedano per esempio: Emanuela Reale, *Valutazione delle pubblicazioni scientifiche e open access* (Reale 2008), Antonella De Robbio, *Analisi citazionale e indicatori bibliometrici nel modello Open Access* (De Robbio 2007), e Mauro Guerrini, *Nuovi strumenti per la valutazione della ricerca scientifica: il movimento dell'open access e gli archivi istituzionali* (Guerrini 2009).

[13] Ad esempio Usage factor, numero di download, ranking, etc.: tutti quei nuovi parametri la cui analisi va sotto il nome di webometria.

Uno degli esperimenti più interessanti è quello prodotto da Plos One che, nel maggio 2009 ha presentato nel suo blog le Article-Level Metrics, un set di metriche applicate al singolo articolo che comprende l'online usage, le citazioni tradizionali nella letteratura scientifica, le citazioni dei blog, dei social bookmarks e i commenti, note e votazioni sull'homepage. Informazioni sul progetto sono disponibili sul blog ufficiale della rivista, EveryONE (http://everyone.plos.org/tag/article-level-

tengano conto delle nuove modalità di trasmissione e di circolazione del sapere rese possibili da Internet.

Prendendo spunto dall'esperienza quotidiana di ogni bibliotecario che lavori a contatto con ricercatori e studiosi, l'articolo, con un taglio assolutamente pragmatico, vuole tentare di fare un po'di chiarezza sui dubbi che a volte si possono presentare durante la ricerca di tali dati.

Non sempre è stato possibile fornire una soluzione certa, o perché non esistente o perché avrebbe richiesto un'analisi più approfondita lontana dal taglio divulgativo-operativo che era stato deciso per questo articolo.

L'Impact factor: il più noto, il più discusso

Partiamo dall'indicatore più noto, ma non per questo il più conosciuto: molti sanno che cosa indica, ma non per tutti è chiaro come viene calcolato e su quale base.

L'Impact factor nasce con lo scopo di indicare il peso (il fattore di impatto") di una rivista all'interno del suo settore disciplinare specifico ed è, matematicamente parlando, il rapporto tra il numero complessivo di citazioni ricevute in un dato anno dagli articoli pubblicati da una certa rivista nei due anni precedenti e il numero di questi ultimi: in pratica il numero medio di citazioni ricevute da un singolo articolo pubblicato da una determinata rivista in un certo lasso di tempo.

Per esempio, se la rivista X ha pubblicato nel biennio 2005-2006 500 articoli e questi hanno ricevuto nel 2007 un totale di 1500 citazioni, il rapporto, cioè l'Impact factor, sarà 1500/500 = 3; il che equivale a dire che, in media, ogni articolo pubblicato nel biennio 2005-2006 nella rivista X ha ricevuto 3 citazioni.

metrics/).

Visto il lasso di tempo necessario per calcolare l'Impact factor, è chiaro che l'edizione più recente del Journal Citation Reports[14] non potrà mai essere quella dell'anno in corso e spesso l'arco temporale è ancora più ampio, causa la grossa mole di informazioni che deve essere raccolta e analizzata per recuperare tale dato[15].

Di conseguenza poiché, quando viene richiesto l'Impact factor, ci si riferisce a quello dell'anno di pubblicazione dell'articolo, se l'articolo è molto recente e tale dato non è stato ancora pubblicato, sarebbe sempre corretto o non riportare il valore o riportare quello indicato sull'ultima edizione del Journal Citation Reports segnalando tra parentesi l'anno dell'edizione consultata.

È bene essere consapevoli che il dato si riferisce alla pubblicazione nel suo insieme, non ad uno specifico articolo, né tantomeno dovrebbe essere assunto come indicatore del valore del singolo autore.

Purtroppo è noto a tutti che, erroneamente, l'Impact factor ha finito per essere interpretato proprio in tal senso con un evidente slittamento concettuale e applicativo, che non può che essere deleterio per la ricerca stessa: l'ottica distorta con cui viene letto tale indicatore spinge infatti gli studiosi a cercare di pubblicare nelle riviste con maggior fattore di impatto", contribuendo così a creare un circolo vizioso della conoscenza, contrario a qualunque concezione di democraticità del sapere" e diffusione dell'informazione[16].

[14]La pubblicazione online che fornisce annualmente, per ognuna delle riviste presenti nella banca dati Web of Science - con l'esclusione della sezione dell'Arts and Humanities Citation Index - oltre all'Impact factor anche il numero totale di citazioni (Total cites), il numero totale di articoli pubblicati (sono escluse alcune tipologia di articoli, come gli editoriali e le lettere, le news e gli abstract dei meeting), l'Immediacy Index, il Cited Half-Life e altri indicatori.

[15]Al momento in cui scriviamo (aprile 2010) l'edizione più recente del Journal Citation Reports è quella del 2008. Normalmente le edizioni escono a giugno dell'anno successivo a quello di riferimento.

[16]Tra i maggiori detrattori dell'utilizzo dell'Impact factor per la valutazione della

La viziosità del circolo è data anche dal fatto che non per tutti i periodici viene calcolato l'Impact factor, perché la banca dati di riferimento sulla base della quale vengono calcolate le citazioni (Web of Science) è in abbonamento e non comprende tutte le riviste esistenti per una data disciplina[17], né alcune particolari tipologie di risorse diverse dall'articolo scientifico che possono aver citato un determinato lavoro.

Il processo di selezione ed inclusione di una rivista in Web of Science e di conseguenza nel Journal Citation Reports si basa inoltre su numerosi criteri [18] che vanno dai requisiti standard di pubblicazione (regolarità e tempestività della pubblicazione, adesione alle convenzioni editoriali internazionali, preferenza per la lingua inglese, utilizzo del peer review) alla copertura internazionale degli autori, dei curatori, del comitato editoriale, ai dati citazionali; criteri che, seppure in gran parte oggettivi e scientificamente validi, possono dare adito ad alcune critiche da parte di chi vede in tali parametri un mezzo per favorire l'oligarchia dei grandi editori scientifici a scapito delle piccole realtà editoriali, non sempre in grado di rispettare gli standard di pubblicazione richiesti.

Di fatto si può considerare l'Impact factor un vero è proprio marchio registrato di proprietà della Thomson Reuters; e se finora non ha avuto rivali, tanto da essere considerato l'indicatore bibliometrico per eccellenza e sinonimo di "valore di un periodico", la situazione

---

ricerca si ricorda Alessandro Figà Talamanca, *L'Impact Factor nella valutazione della ricerca e nello sviluppo dell'editoria scientifica*, intervento al IV seminario SINM 2000 (Figa' Talamanca 2000): un modello di sistema informativo nazionale per aree disciplinari. Cfr. anche Per O Seglen, *Why the impact factor of journals should not be used for evalutating research* (O Seglen 1997).

[17] Attualmente il Science Citation Index Expanded, una delle 7 banche dati che compongono Web of Science, indicizza 6650 riviste, mentre il Social Science Citation Index ne indicizza 1950.

[18] http://thomsonreuters.com/products_services/science/free/essays/journal_selection_process/.

attualmente sta cambiando grazie all'evoluzione dei processi editoriali e alla nascita di parametri alternativi e ad esso paragonabili, come ad esempio Scimago e l'Eigenfactor.

Scimago e Eigenfactor: l'alba di una nuova era?

Il portale Scimago[19] nasce nel 2007 da un gruppo di ricerca del Consejo Superior de Investigaciones Científicas (CSIC), dell'Università di Granada, dell'Estremadura, dell'Università Carlos III di Madrid e dell'Università Alcalá de Henares e mira a candidarsi come il più diretto avversario del Journal Citation Reports. La base dati utilizzata da Scimago per calcolare il fattore d'impatto" delle riviste, noto come Scimago Journal Rank Indicator[20] (SJR), è Scopus[21] della Elsevier, una banca dati in abbonamento che si sta ponendo come la maggiore rivale di Web of Science della Thomson Reuters.

Contrariamente all'Impact factor, la cui visualizzazione è vincolata ad un abbonamento al Journal Citation Reports, Scimago, pur avvalendosi di informazioni che sono disponibili solo a pagamento,

[19] http://www.scimagojr.com/index.php.

[20] Lo Scimago Journal Rank Indicator (SJR) viene definito dal glossario in linea del portale «a measure of journal's impact, influence or prestige. It expresses the average number of weighted citations received in the selected year by the documents published in the journal in the three previous years». Per un'analisi della nascita del SJR si veda Borja Gonzalez-Pereira, Vicente P. Guerrero-Boteb e Félix Moya-Anegónc, *The SJR indicator: A new indicator of journals' scientific prestige* (Gonzalez-Pereiraa, Guerrero-Boteb e Moya-Anegónc 2009).

[21] Attualmente Scopus indicizza circa 18000 risorse (16500 riviste peer reviewed). Le politiche di selezione e inclusione di una rivista nella banca dati sono presentate nelle FAQ (http://info.scopus.com/scopus-training/faqs/).

garantisce un accesso gratuito al proprio Journal rank indicator[22], basato sul modello di ranking di Google.

Così come nel Journal Citation Reports, è possibile ricercare per singola rivista o per disciplina, ma Scimago permette anche di scoprire che posizione ricopre un determinato paese nella ricerca in uno specifico ambito disciplinare o in una sua più ristretta categoria, tramite le opzioni di Country ranking e Country Search[23].

Colpisce il fatto che il portale dia la possibilità di pubblicare su un proprio sito le tabelle di indicatori delle singole riviste, semplicemente copiando il codice html necessario a creare il link.

È evidente come, sotto a tutto il progetto Scimago, ci sia una chiara scelta culturale che si rifà al mondo dell'accesso aperto: ogni informazione è gratuita e visibile a tutti, esportabile e condivisibile.

Pesa però, in una misura al momento non definibile, sulle possibilità di diffusione del SJR come indicatore alternativo all'Impact factor lo stretto e fondamentale legame con una banca dati a pagamento quale Scopus, la cui palese volontà di porsi come l'attuale unica vera rivale di Web of Science fa sì che, per potersi realmente confrontare con la "corazzata" dell'ISI, Scopus tenda a replicarne il modello commerciale difficilmente accostabile ad un progetto Open access.

Il rischio è che, nel tentativo di sottrarre ampie fette di mercato all'ISI, Elsevier finisca per imporre un obbligo di abbonamento anche per l'accesso ai dati citazionali che poggiano su Scopus, blindando

[22]Oltre al Journal rank indicator, Scimago indica annualmente per ogni rivista anche: l'H-index, il numero di articoli pubblicati, il numero di referenze bibliografiche, il numero di citazioni ricevute dagli articoli pubblicati nella rivista nei 3 anni precedenti, le autocitazioni, i documenti citabili (articoli, revisioni e conference papers), etc. Per una breve introduzione ai vari indicatori forniti da Scimago si veda Grupo Scimago, *SCImago journal & country rank: un nuevo portal, dos nuevos rankings* (Scimago 2007).

[23]Una possibilità simile è offerta anche dal National Science Indicators, strumento a pagamento della Thomson Reuters.

anch'essa il recupero delle informazioni all'interno di un'offerta commerciale a pagamento.

Si pone sulla stessa linea di libera accessibilità e gratuità al momento garantita da Scimago anche il progetto Eigenfactor[24], nato come progetto di ricerca accademica non-commerciale sponsorizzato dal Bergstrom lab al Department of Biology dell'Università di Washington. Così come Scimago anche l'Eigenfactor si rifà all'algoritmo di ranking di Google per creare le liste in ordine di peso" delle riviste nelle varie discipline.

Definito dai suoi creatori

> «a measure of the overall value provided by all of the articles published in a given journal in a year»

[25] ha come caratteristica principale la volontà di tener conto dei differenti standard citazionali tra una scienza e l'altra per calcolare l'impatto delle riviste; cerca, cioè, di equilibrare le differenze citazionali in modo da poter effettuare una comparazione più contestualizzata tra le varie aree disciplinari.

L'attenzione alle peculiarità delle singole materie sta anche nell'aver esteso il riferimento temporale delle citazioni da due a cinque anni, perché non tutte le aree scientifiche hanno la stessa tempistica citazionale[26].

---

[24] http://www.eigenfactor.org/index.php . Oltre all'Eigenfactor viene calcolato anche un altro indicatore, l'Article influence e colpisce il fatto che proprio quest'ultimo sia definito dai responsabili del progetto come «*a measure of a journal's prestige based on per article citations and comparable to Impact Factor*» [corsivo mio]. Per una descrizione dettagliata si veda Carl. T. Bergstrom, *Eigenfactor: Measuring the value and prestige of scholarly journals* (Bergstrom 2007), e Carl. T. Bergstrom, Jevin D. West e Marc A. Wiseman, *The Eigenfactor Metrics* (Bergstrom, West e Wiseman 2008).

[25] Questa è la definizione che viene fornita, come glossario, nella pagina dei risultati dopo aver effettuato una ricerca.

[26] Il progetto sintetizza i suoi punti di forza in una lista di 6 voci disponibile alla pagina http://www.eigenfactor.org/whyeigenfactor.htm. Un'altra differenza con

L'attuale limite dell'Eigenfactor è però costituito dal fatto che, come ci dicono gli stessi autori del progetto nella pagina delle FAQ[27], la base dati utilizzata come fonte delle informazioni è proprio il Journal Citation Reports[28].

Si può facilmente dedurre che, poggiando sia l'Impact factor sia lo SJR sia l'Eigenfactor su banche dati commerciali a pagamento, l'unico vero discrimine[29] tra i tre indicatori bibliometrici è dato dall'attuale libertà di accesso garantita da Scimago e Eigenfactor contro la consultazione in solo abbonamento del Journal Citation Reports.

Non spetta pertanto a noi decidere quale dei vari strumenti debba eventualmente prevalere[30], ma è compito della nostra professione

l'Impact factor da ricordare è che l'Eigenfactor non tiene conto delle autocitazioni tra articoli pubblicati nella stessa rivista.

[27]«5. Where do you get your data? The citation data used at Eigenfactor come from Thomson Scientific's Journal Citation Reports (JCR). At present, we have full data only for those journals listed in the JCR. The JCR provides detailed information about the nearly 7,000 journals in the sciences and social sciences which are indexed in Thomson's Web of Science citation database. Journals in the sciences and social sciences. In addition, we are able to assign Eigenfactor scores to more than 100,000 reference items that are cited by the JCR-listed journals. 7. Why isn't my journal ranked by Eigenfactor™ score? How can I get it listed? The journal citation data for the Eigenfactor website come from Thompson Scientific's Journal Citation Reports (JCR) for each year 1995-2006. Thus, to be listed on the Eigenfactor website, a journal either must be listed in the JCR or, minimally, be cited in the last 5 years by some journal that is listed in the JCR. At present, we have no other means for facilitating inclusion in the Eigenfactor database.» Vedi http://www.eigenfactor.org/faq.htm.

[28]Non è un caso, infatti, che dal 2007 sia l'Eigenfactor che l'Article influence siano reperibili anche nel Journal Citation Reports.

[29]Anche se esistono delle differenze tra i vari algoritmi di ricerca che stanno alle base dei tre indicatori bibliometrici analizzati, in linea generale si può parlare di sostanziale convergenza e somiglianza. Cfr. Ronald Rousseau e Stimulate 8 Group, *On the relation between the WoS impact factor, the Eigenfactor, the SCImago Journal Rank, the Article Influence Score and the journal h-index* (Rousseau e Group 2009).

[30]Anche perché, come fa notare giustamente Moed, «there is no single perfect" indicator of journal performance. The scholarly communication system is highly

indicare le alternative esistenti e promuoverne l'uso corretto laddove non sia esplicitamente indicata l'obbligatorietà di utilizzo di una fonte sull'altra.

Citation index: il numero delle citazioni come fulcro di tutto

Siamo partiti dall'Impact factor per il nostro viaggio tra le pieghe degli indicatori bibliometrici, ma il vero punto di partenza, la base su cui tutto poggia è uno solo: il numero di citazioni che un articolo, e di conseguenza a livello più ampio un autore o una rivista, ha ricevuto.

Questo è il dato essenziale senza il quale non potrebbero essere calcolati i vari indicatori.

Il numero di citazioni (conosciuto come Citation index[31]) si spiega concettualmente da sé, ma, contrariamente all'Impact factor, non si può considerare come un marchio registrato, un dato univoco

---

complex, citations constitute one of its representations – though a most valid and useful one – and journal performance is a multi-dimensional concept that cannot be expressed in any single measure. The adequacy of a journal impact measure is related to the type of use made of it, and the type of research question addressed. A particular indicator may be appropriate in one context, and less appropriate in another.» (Moed 2010).

[31]Purtroppo non sempre l'accuratezza terminologica è garantita: per esempio la Regione Toscana nel suo bando per il Programma per la Ricerca Regionale in Materia di Salute 2009 (http://www.regione.toscana.it/regione/multimedia/RT/documents/2009/07/20/0db3fc4c7083b8dcce0d6850f58f5b49_allegatoabando.pdf) chiedeva ai candidati di fornire «il numero delle citazioni ottenute nell'ultimo triennio dal 2006 (*Quotation Index*) [corsivo mio] secondo quanto riportato dall'ISI Web of Knowledge», assegnando ai termini to cite" e to quote" una sinonimia che, se valida per uso generale, non lo è per uno specifico laddove il significato semantico ha particolare peso.

vincolato, perché il suo valore può cambiare secondo la banca dati di riferimento.

Se è vero infatti che, in linea teorica, il numero di citazioni ricevute da un articolo, non può che essere unico, vari sono in realtà i fattori che contribuiscono alla sua variazione: ad esempio, la copertura delle singole banche dati sia come numero di riviste indicizzate, sia come anni di pubblicazione o di sottoscrizione; oppure i criteri di spoglio e gli algoritmi di ricerca, per cui alcune tipologie di articoli possono non essere incluse e quindi non essere conteggiate.

Un corretto recupero di tale dato dovrebbe attualmente, in mancanza di specificazioni chiare e, se possibile, essere effettuato almeno su tre banche dati diverse: Web of Science, Scopus e Google Scholar.

Se i risultati differiscono dovrebbero essere riportati tutti e tre con l'indicazione delle basi dati su cui sono stati calcolati.

Il recupero del Citation index su Web of Science deve tenere conto di alcuni aspetti propri di questa banca dati. Web of Science dà, infatti, la possibilità di ricercare sia il numero di citazioni per il singolo articolo, sia il numero di citazioni complessivo per il singolo autore o singola rivista (con l'opzione Cited Reference Search).

La scelta di utilizzare l'una o l'altra modalità di ricerca dipende dai motivi per cui si sta cercando tale dato: la Cited Reference Search, per esempio, non mostra gli articoli che non hanno ricevuto nessuna citazione, mentre ad un autore potrebbe interessare sapere quali delle sue pubblicazioni non sono mai state citate, quindi quale filone di ricerca potenziare o su quale argomento concentrare le proprie energie.

Nella Cited Reference Search è attiva inoltre un'applicazione (Author Finder) che dovrebbe servire a risolvere uno dei maggiori problemi che deve spesso fronteggiare chi si trova a lavorare con questi strumenti: le omonimie.

Author Finder aiuta, con una serie di step predefiniti, a disambi-

guare gli autori omonimi, ma il suo utilizzo non è garanzia di certezza. Il problema è particolarmente evidente quando si arriva alla pagina relativa all'affiliazione di un autore ad un ente: quest'ultimo spesso è indicato con numerose varianti, non sempre tutte immediatamente raggruppabili, senza considerare il fatto che un autore, nella sua vita scientifica, può aver più volte cambiato affiliazione.

Può essere visto come un tentativo di risolvere il problema delle omonimie anche ResearcherID[32], una comunità scientifica promossa dalla Thomson Reuters che permette ai ricercatori di registrarsi gratuitamente, ricevendo in cambio un numero identificativo univoco, e di creare un proprio profilo aggiornabile in ogni istante: l'associazione di un descrizione costruita dall'autore stesso con le proprie pubblicazioni, le varie affiliazioni, le proprie attività professionali e un numero identificativo univoco dovrebbe contribuire a disambiguare le omonimie.

La lista di pubblicazioni che ogni autore inserisce nel proprio profilo viene sottoposta ad analisi citazionale incrociandola con la banca dati Web of Science e fornendo quindi il numero complessivo di citazioni, la media di citazioni per articolo e l'h-index; ne consegue che, pur avendo l'autore la possibilità di inserire nella lista tutte le sue pubblicazioni, solo gli articoli che sono stati indicizzati in Web of Science vengono utilizzati per il calcolo citazionale.

Le potenzialità di tale prodotto stanno nella possibilità di interrogazione gratuita da parte di chiunque[33], non solo degli utenti registrati, il che significherebbe la possibilità di recuperare importanti indicatori bibliometrici svincolandola da un abbonamento a Web of Science.

Si tratta dello stesso rapporto che abbiamo visto precedentemente tra Scimago e Scopus, in questo caso però applicato non alla

[32]http://www.researcherid.com/Home.action.

[33]La visualizzazione delle informazioni è garantita però solamente per quei profili che gli autori hanno settato come "pubblici", cioè visibili a tutti.

rivista ma al singolo autore. Il limite dello strumento sta, ad oggi, proprio nella volontarietà dell'iscrizione alla comunità e nella partecipazione attiva dell'utente registrato, la cui tempestività o meno nell'aggiornamento delle informazioni che lo riguardano, in particolare della propria bibliografia, influisce sull'attendibilità degli indicatori recuperati.

Scollegato da una partecipazione attiva dell'autore alla propria disambiguazione è invece Author Identifier, lo strumento di Scopus che si attiva con le opzioni di ricerca Author Search e Affiliation Search: accoppiando e de-duplicando automaticamente i nominativi degli autori, assegna un numero unico di identificazione a tutti coloro che hanno pubblicato articoli su riviste indicizzate dalla banca dati. Per aumentare la percentuale di precisione l'algoritmo utilizza anche dati supplementari associati agli articoli dell'autore, quali affiliazione, cronologia delle pubblicazioni, titolo della fonte, area tematica e coautori.

Nel momento in cui l'utente svolge una ricerca per autore, l'algoritmo che sta alla base di Author Identifier genera una lista di risultati con le possibili varianti di ogni singolo nome e i dati associati (affiliazione, area disciplinare di riferimento, numero totale di documenti associati al singolo autore e il link ad una pagina con i dettagli sull'autore).

Se l'identificazione è certa e la disambiguazione efficace, la pagina con i dettagli dell'autore (Author details) fornisce due indicatori citazionali importanti: il Citation index totale e l'H-index.

La visualizzazione del numero di citazioni per singolo articolo o la possibilità di intervenire nella selezione dei documenti recuperati (nel caso in cui l'identificazione e disambiguazione dell'autore non convincano) è affidata alla Citation Overview attivabile tramite la funzione di Citation tracker.

Lo stesso problema della disambiguazione degli omonimi si

ritrova anche su Google Scholar[34].

Tra i numerosi software open source esistenti in rete per il recupero di indicatori citazionali utilizzando come banca dati Google Scholar, il più diffuso è sicuramente PoP (Publish or Perish)[35], sviluppato dalla professoressa Anne-Will Harzing dell'Università di Melbourne in Australia.

Fornisce moltissimi indicatori bibliometrici[36] sia per gli autori che per le riviste, ma non permette una disambiguazione efficace delle omonimie. Il sistema consente, attraverso l'utilizzo degli operatori booleani, di inserire le possibili varianti note della forma del nome di un autore, di escluderne un altro, ha dei filtri disciplinari, ma di fatto la risoluzione delle omonimie è possibile solo a posteriori sui risultati della ricerca, selezionando e deselezionando dall'elenco ottenuto le voci di interesse con un aggiornamento automatico e immediato degli indicatori citazionali associati.

H-index: l'ultimo arrivato nella grande famiglia

L'ultimo tra gli indicatori bibliometrici ad essere stato sviluppato è l'H-Index (o Indice di Hirsch), messo a punto nel 2005 da Jorge E. Hirsch (Hirsch 2005), professore all'Università della California di San Diego, nel tentativo di trovare un indicatore che equilibrasse e

[34]Vasta banca dati di letteratura accademica a consultazione gratuita. Contiene documenti documenti approvati per la pubblicazione, tesi, libri, abstract e articoli di case editrici accademiche, ordini professionali, database di studi non ancora pubblicati, università e altre organizzazioni accademiche.

[35]Liberamente scaricabile dall'indirizzo http://www.harzing.com/pop.htm.

[36]Oltre a semplici dati statistici (numero di documenti recuperati, media di articoli per anno, media di autori per articolo, media di articoli per autore, etc.) fornisce anche: h-index, g-index, hc-index, hI-index, hI-norm., AWCR, AW-index, AWCRpA, e-index, hm-index (http://www.harzing.com/pop.htm#metrics).

rendesse comparabili autori molto profilici, ma con un basso numero di citazioni per articolo (ricca produzione, ma di scarso rilievo) e studiosi meno produttivi, ma con un alto rapporto citazione/articolo (produzione numericamente scarsa, ma di grande impatto).

Definito come un valore h pari al numero di articoli di un autore che hanno ricevuto almeno h citazioni (h-index 3 significa che un autore ha scritto almeno tre articoli che hanno ricevuto almeno tre citazioni), è solitamente riferito alla produzione della singola persona, ma può essere applicato anche ad un periodico, come abbiamo visto con Scimago.

Come per il Citation index, il suo recupero, proprio in virtù dello stretto legame che li unisce, può essere eseguito nelle banche dati già citate, Web of Science[37], Scopus e Google Scholar, con il conseguente rischio di ottenere risultati diversi a seconda della fonte utilizzata.

Sebbene sviluppato da pochissimi anni è diventato rapidamente uno degli indicatori più noti e maggiormente usati ai fini della valutazione; ma, nonostante tanto successo, anche l'H-index, così come gli altri indici precedentemente illustrati, non è privo di limiti.

Evidente per esempio è la sua debolezza di fronte ad autori che, seppur riconosciuti come autorevoli dalla propria comunità scientifica di riferimento, hanno avuto una carriera molto breve: si pensi ad un personaggio come Evariste Galois[38], morto a soli ventuno anni, la cui unica opera venne pubblicata nel numero di ottobre-novembre del 1843 del Giornale di matematica pura e applicata.

Basandosi unicamente sul dato numerico del Citation index di un articolo, l'H-index non tiene conto inoltre di tre fenomeni importanti:

[37]In Web of Science il dato è parzialmente nascosto, perché visualizzabile solo attraverso la funzione Citation Report che, contrariamente a quanto la logica farebbe supporre, non si attiva nella Cited Reference Search, ma dalla lista dei risultati ottenuti dalla ricerca per autore effettuata con Search.

[38]Matematico (1811-1832), a lui si deve la teoria che porta il suo nome e che costituisce un'importante branca dell'algebra astratta.

le autocitazioni che, considerando l'intero corpus produttivo di un autore, possono essere molte e accrescere artificialmente l'indice; il contesto delle citazioni e la co-produzione degli articoli.

Soprattutto nel campo delle scienze dure, gli articoli sono scritti a numerose mani e l'ordine con cui gli autori si presentano riflette il peso del loro contributo alla ricerca: primo e ultimo autore hanno un ruolo più influente nel progetto rispetto a coloro che sono elencati nel mezzo.

Considerare il numero di citazioni ricevute da un articolo non tenendo conto di tale realtà di fatto costituisce un' involontaria falsificazione[39]

La sua applicazione più comune, che considera la produzione di uno studioso nel suo complesso[40], mira a rappresentare con un secco dato numerico, inevitabilmente semplificando una realtà ben più complessa, l'impatto complessivo del lavoro scientifico di un autore; ma in un'ottica più funzionale alla valutazione della ricerca come strumento di analisi economica, può essere applicato anche ad un arco temporale specifico per valutare la validità di un investimento finanziario fatto o da fare su una determinata persona o organismo.

[39]Proprio nel tentativo di risolvere tale anomalia sì è cercato di raffinare l'indice di Hirsch, proponendone delle varianti.(Schreiber 2008; Wan, Hua e Rousseau 2007). Si muove in questo senso anche Publish or Perish su Google Scholar, offrendo anche il Normalized Individual h-index (hI,norm): il software divide il numero di citazioni di un articolo per il numero di autori che vi hanno contribuito, normalizzando così il dato, poi sul numero ottenuto calcolano l'H-index.

Per una chiara sintesi sulle varianti sviluppate a partire dall'indice di Hirsch si veda Sergio Alonso, Francisco Javier Cabrerizo, Enrique Herrera Viedma e Francisco Herrera, *H-index: A Review Focused in its Variants, Computation and Standardization for Different Scientific Fields* (Alonso *et al.* 2009).

[40]È sempre bene rimarcare che anche il concetto di interezza" non è mai assoluto, perché le varie banche dati analizzate non raccolgono la reale intera produzione di un autore, ma una - forse- inevitabile selezione basata sui canali di distribuzione, sugli anni di copertura, etc.

Per esempio, un'università potrebbe voler valutare la ripartizione delle risorse tra i vari dipartimenti verificando anche l'h-index del personale afferente alle singole sedi per i soli lavori scientifici prodotti e pubblicati nell'arco di tempo di appartenenza alla specifica struttura; in tal caso il calcolo dell'H-index verrebbe applicato solamente a una porzione dell'intera produzione di uno studioso.

La più ovvia riflessione che si possa fare, a conclusione di questo rapido excursus nel magmatico mondo degli indicatori bibliometrici, è che solo una conoscenza approfondita di tali strumenti può permetterne un uso consapevole.

Il ruolo attivo di consulenza e partecipazione che le biblioteche possono giocare nella costruzione di un sistema di valutazione della ricerca che sia realmente efficace passa anche da qui[41].

Le biblioteche accademiche o scientifiche potrebbero mettere infatti a disposizione le proprie competenze a più livelli [42]: lavorando a stretto contatto con i Nuclei di valutazione della ricerca, presenti in numerosi Atenei, occupandosi, ad esempio, della raccolta dei dati, ma anche – e soprattutto – estendendo il sostegno all'attività di ricerca di docenti e ricercatori anche alle fasi successive alla pubblicazione vera e propria dei risultati. Le biblioteche, infatti, attualmente svolgono la loro attività di supporto alla ricerca mettendo a disposizione le risorse informative necessarie, garantendo un catalogo in linea ordinato e coerente, accessi alle risorse elettroniche solidi e

[41]Sul ruolo delle biblioteche nelle valutazione della ricerca si veda il recente report pubblicato da OCLC redatto da John MacColl, *Research Assessment and the Role of the Library. Report produced by OCLC Research* (MacColl 2010).

Anche i sistemi bibliotecari degli atenei italiani rimarcano la necessità di rafforzare il proprio ruolo all'interno dei processi valutativi della ricerca, come si legge in Elena Giglia, *I sistemi bibliotecari di Ateneo: realtà a confronto e prospettive future, Roma, 18 febbraio 2010 - Appunti per una sintesi* (Giglia 2010).

[42]Molto interessante a tal proposito l'analisi, seppure rispondente alla realtà statunitense, presentata da Herther («Research evaluation and citation analysis: key issues and implications») in particolare p. 367-369

controllati, svolgendo servizi di prestito interbibliotecario e di fornitura dei documenti per ovviare alle inevitabili lacune del proprio posseduto documentario: attività ormai ampiamente consolidate. Raramente però il personale bibliotecario viene coinvolto nelle fasi successive della ricerca: la scelta del miglior canale di pubblicazione (o di canali alternativi di pubblicazione), la valutazione dell'impatto che i risultati della ricerca hanno avuto nelle varie comunità di riferimento o la constatazione del peso che il proprio operato scientifico ha avuto nella ripartizione dei fondi all'interno dell'ente di appartenenza. L'acquisizione delle competenze atte a svolgere queste valutazioni e l'apprendimento degli strumenti e degli indicatori necessari avvengono spesso in maniera disorganica, lasciate alla buona volontà del singolo docente o ricercatore. I bibliotecari, invece, potrebbero inserirsi in questo momento di formazione, o fornendo all'utente strumenti informativi per l'auto-formazione, come tutorial, guide on line, slides esplicative, oppure predisponendo attività di consulenza e di reperimento dei dati in forma strutturata: recupero degli indicatori bibliometrici dietro invio tramite e-mail della bibliografia aggiornata, informazioni via telefono sulle riviste con maggior impact factor in una specifica area disciplinare per la scelta della rivista cui inviare un articolo, etc. La volontà di rendere autonomi gli utenti non dovrebbe infatti impedire il soddisfacimento di bisogni precisi e puntuali del singolo, in un'ottica di ottimizzazione dei tempi e delle risorse. Accanto a queste attività, inoltre, le biblioteche potrebbero integrare i propri corsi sull'utilizzo delle risorse informative con appositi moduli relativi alla valutazione della ricerca, per quelli che sono gli aspetti di propria pertinenza, rivolti principalmente a laureandi, specializzandi o dottorandi; coloro, cioè, che potenzialmente si occuperanno di ricerca in futuro. In fondo non si tratterebbe altro che di strutturare e integrare attività e competenze già previste nella realtà bibliotecaria: come spesso succede,

solo una questione di integrazione e organizzazione.

Riferimenti bibliografici

ALONSO, SERGIO *et al.* (2009), «H-index: A Review Focused in its Variants, Computation and Standardization for Different Scientific Fields», *Journal of Informetrics*, 3, 4, pp. 273–289, <http://eprints.rclis.org/16434/>.

BERGSTROM, CARL T. (2007), «Eigenfactor: Measuring the value and prestige of scholarly journals», *C&RL News*, 68, 52, pp. 314–316, http://crln.acrl.org/content/68/5/314.full.pdf.

BERGSTROM, CARL T., JEVIN D. WEST e MARC A. WISEMAN (2008), «The Eigenfactor Metrics», *Journal of Neuroscience*, 28, 45, pp. 11433–11434, http://www.eigenfactor.org/BergstromEtAl08.pdf.

CIVR (2006), *Linee guida per la valutazione della ricerca*, Ministero dell'Istruzione, dell'Università e della Ricerca, http://vtr2006.cineca.it/documenti/linee_guida.pdf.

DE ROBBIO, ANTONELLA (2007), «Analisi citazionale e indicatori bibliometrici nel modello Open Access», *Bollettino AIB*, 47, 3, pp. 257–287, http://eprints.rclis.org/11999/.

— (2010), *L'Open Access come strategia per la valutazione delle produzioni intellettuali*, in *CIBER 1999-2009*, a cura di Paola Gargiulo e Domenico Bogliolo, Ledizioni, pp. 104–124, http://eprints.rclis.org/17742/.

DI DONATO, FRANCESCA (2010), *La scienza e la rete*, Firenze University Press, http://www.fupress.com/Archivio/pdf/3867.pdf.

FIGA' TALAMANCA, ALESSANDRO (2000), «L'Impact Factor nella valutazione della ricerca e nello sviluppo dell'editoria scientifica», Intervento al IV Seminario SINM, http://siba2.unile.it/sinm/4sinm/interventi/fig-talam.htm.

GARFIELD, EUGENE (1955), «Citation Indexes for Science: a new dimension in documentation through association of ideas», *Science*, 122, pp. 108–111.

— (2006), «The history and meaning of the journal impact factor», *JAMA*, 295, 1, pp. 90–93, http://www.garfield.library.upenn.edu/papers/jamajif2006.pdf.

GARFIELD, EUGENE, HIRVING H. SHER e RICHARD J. TORPIE (1964), *The use of citation data in writing the history of science*, Air Force Office of Scientific Research, http://www.garfield.library.upenn.edu/papers/useofcitdatawritinghistofsci.pdf.

GIGLIA, ELENA (2010), «I sistemi bibliotecari di Ateneo: realtà a confronto e prospettive future, Roma, 18 febbraio 2010 - Appunti per una sintesi», *AIDAinformazioni*, 28, 1-2, http://www.aidainformazioni.it/pub/giglia(2)122010.html.

GONZALEZ-PEREIRAA, BORJA, VICENTE P. GUERRERO-BOTEB e FÉLIX MOYA-ANEGÓNC (2009), «The SJR indicator: A new indicator of journals' scientific prestige», Avalaibleathttp://arxiv.org/ftp/arxiv/papers/0912/0912.4141.pdf.

GUERRINI, MAURO (2009), «Nuovi strumenti per la valutazione della ricerca scientifica: il movimento dell'open access e gli archivi istituzionali», *Biblioteche oggi*, 27, 8, pp. 7–17, http://eprints.rclis.org/17066/.

HERTHER, NANCY K., «Research evaluation and citation analysis: key issues and implications», *The Electronic Library*, 27.

HIRSCH, JORGE E. (2005), «An index to quantify an individual's scientific research output», *Proceedings of the National Academy of Sciences of the United States of America*, 102, 46, pp. 16569–16572.

MACCOLL, JOHN (2010), *Research Assessment and the Role of the Library. Report produced by OCLC Research*, OCLC, http://www.oclc.org/research/publications/library/2010/2010-01.pdf.

MOED, HENK F. (2010), «Measuring contextual citation impact of scientific journals», Article in press.

O SEGLEN, PER (1997), «Why the impact factor of journals should not be used for evalutating research», *British Medical Journal*, 314, 7079, pp. 497–502.

REALE, EMANUELA (2008), «Valutazione delle pubblicazioni scientifiche e open access», http://eprints.rclis.org/13493/.

ROUSSEAU, RONALD e STIMULATE 8 GROUP (2009), «On the relation between the WoS impact factor, the Eigenfactor, the SCImago Journal Rank, the Article Influence Score and the journal h-index», http://eprints.rclis.org/16448/.

SCHREIBER, MICHAEL (2008), «A modification of the h-index: The hm-index accounts for multi-authored manuscripts», *Journal of Informetrics*, 2, 3, pp. 211–216.

SCIMAGO, GRUPO (2007), «SCImago journal & country rank: un nuevo portal, dos nuevos rankings», *El profesional de la Información*, 16, 6, pp. 645–646, http://www.elprofesionaldelainformacion.com/contenidos/2007/noviembre/11.pdf.

WAN, JIN-KUN, PING-HUAN HUA e RONALD ROUSSEAU (2007), «The pure h-index: calculating an author's h- index by taking co-authors into account», *Collnet Journal of Scientometrics and Information Management*, 1, 2, pp. 1–5, http://eprints.rclis.org/11401/.

Informazioni

L'autore

Tessa Piazzini
Università degli studi di Firenze. Biblioteca Biomedica.
Email: tessa.piazzini@unifi.it

Il saggio

Data di submission: 2010-04-18

Data di accettazione: 2010-04-20

Ultima verifica dei link: 2010-05-22

Data di pubblicazione: 2010-06-15